PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/45, C07K 14/115, A61K 39/165, G01N 33/50, C07K 16/10, C12N 9/24

A2

(11) International Publication Number: WO 99/02695

(43) International Publication Date: 21 January 1999 (21.01.99)

(21) International Application Number: PCT/NL98/00390

(22) International Filing Date: 8 July 1998 (08.07.98)

(30) Priority Data:
97202100.0 8 July 1997 (08.07.97) EP

(71) Applicant *(for all designated States except US)*: STICHTING INSTITUUT VOOR DIERHOUDERIJ EN DIERGEZONDHEID [NL/NL]; Edelhertweg 15, NL–8219 PH Lelystad (NL).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LANGEDIJK, Johannes, Petrus, Maria [NL/NL]; Spuistraat 64G, NL–1012 TW Amsterdam (NL). VAN OIRSCHOT, Johannes, Theodorus [NL/NL]; Oostrandpark 18, NL–8212 AN Lelystad (NL).

(74) Agent: SMULDERS, Th., A., H., J.; Vereenigde Octrooibureaux, Nieuwe Parklaan 97, NL–2587 BN The Hague (NL).

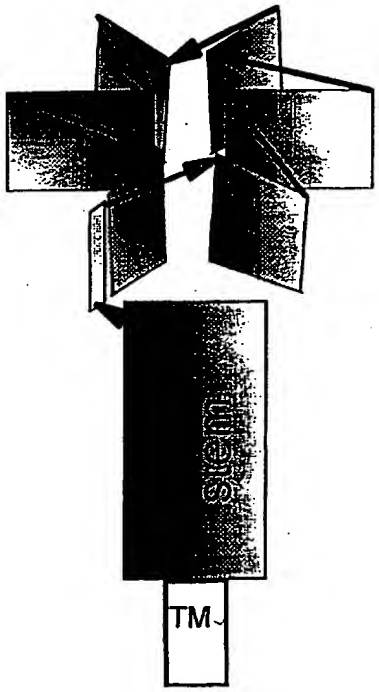
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: EPITOPES AND ACTIVE SITES OF PARAMYXOVIRIDAE PROTEINS AND USES THEREOF

(57) Abstract

The invention relates to the field of paramyxoviridae, vaccines against infections by such viruses, diagnostics for detecting such viruses and targets for therapeutics against such viruses. In particular, the invention relates to 3–D models identifying a proteinaceous substance comprising at least one virus epitope derived from the attachment protein of a virus from the family of paramyxoviridae, said epitope corresponding to an antigenic site present on the HN protein of paramyxovirus, which site is identified as one of loop β1L01, β1L23, β2L01, β2L23, β3L01, β3L23, β4L01, β4L23, β5L01, β5L23, β6L01 and β6L23, or a functional equivalent thereof. Also, the invention relates to a substance blocking the enzymatic activity of the morbillivirus H protein.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

EPITOPES AND ACTIVE SITES OF PARAMYXOVIRIDAE PROTEINS AND USES THEREOF

The present invention relates to the fields of molecular biology and medicinal and/or diagnostic products designed through molecular biology and molecular modelling. In particular the present invention relates to the field of paramyxoviridae, vaccines against infections by such viruses, diagnostics for detecting such viruses and therapeutics against such viruses.

Paramyxoviridae are single stranded RNA viruses of which the genome is a negative RNA strand. This means that the viral RNA cannot be directly translated into viral proteins by the infected host cell. Only the complementary strand of the genomic RNA can be translated. Eukaryotic cells which are the target for these viruses do not possess enzymes that are able to transcribe RNA from an RNA template. Thus the virus has to provide these transcription enzymes itself. The virus does this by introducing its RNA in the cell as a ribonucleoprotein complex. This complex comprises the genomic (-) RNA, together with three proteins, i.e. the nucleoprotein (NP), the phosphoprotein (P) and the large protein (L), the latter two of which form the RNA dependent RNA polymerase necessary for transcription of the viral genome.

The family *Paramyxoviridae* contains two subfamilies: the *Pneumovirinae* and *Paraflumorbillivirinae.* The *Paraflumorbillivirinae* can be subdivided in three genera: *Respirovirus* (sendai, parainfluenza type I, parainfluenza type III); *Rubulavirus* (mumps, simian type 5, Newcastle disease, parainfluenza type II and parainfluenza type IV) and

Morbillivirus (measles, rinderpest and the distemper viruses PDV and CDV). *Respirovirus* and *Rubulavirus* are often combined in one genus: *paramyxovirus*. The *Pneumovirinae* are classified as a separate genus because of differences in the diameter of the nucleocapsid and the lack of detectable hemagglutination and neuraminidase activity (27, 49). They also differ in aspects of viral RNA and protein structure (8). It may be clear from the above that this group of viruses includes a number of important pathogens, both for humans and animals. It would of course be very useful if antiviral agents with high specificity for these viruses could be developed. This is one of the objects of the present invention. It would of course also be useful if vaccines could be developed against these viruses, especially vaccines which elicit a protective response in the vaccinated subject, which response can however be distinguished from the response to an infection of a wild type virus. In order to be able to attain the above objectives, the present inventors studied the structure of a number of paramyxoviridae in depth. In particular the structural proteins possibly involved in mechanisms of infection were studied. Earlier structural studies comparing influenza virus N and paramyxovirus HN (Jorgensen et al 1987; Colman et al, 1993) failed in identifying a precise location of relevant sites, postulating incorrect alignments, and giving only a rough estimate of sheets composing a globular head of a protein.

The *Paramyxoviridae* are enveloped viruses that contain two envelope glycoproteins, the fusion protein (F) and the attachment protein (HN, H or G). The attachment protein HN of *Paramyxoviruses* contains both hemagglutination and neuraminidase (sialidase) activity, like influenzavirus neuraminidase (N), and binds and cleaves terminal sialic acids. The attachment protein (H) of *Morbilliviruses* has hemagglutinin activity, but neuraminidase activity has never been described. Both H and HN are globular proteins of the same

size and the position of these attachment proteins in the genome organization is conserved. The function of the neuraminidase activity of viruses is not well understood. It has been shown that the *orthomyxovirus* influenza virus N protein is necessary to facilitate the release of progeny virus from infected cells (47). Cleavage of sialic acids releases the virus from the glycosylated cellular membrane proteins. Another possible role of the neuraminidase may be the transport of the virus through the sialic acid-rich mucus-layer that protects internal body parts from harmful agents which will be discussed herein. It has been demonstrated for several paramyxoviridae that HN is necessary for the initial fusion. It has been proposed that both F and HN act in concert to establish infection however, the requirement for HN for this process is still questioned (reviewed in 28). Furthermore, a type-specific functional interaction between F and HN of some *paramyxoviruses* is required (3, 13, 20, 52). Similarly, a specific interaction is proposed for F and H of a *morbillivirus* (5).

The invention discloses, based on the multiple sequence alignment of a diverse set of neuraminidases of varying origin, 3-D models for *paramyxoviridae* HN and H. We compared the sequence and structure of morbillivirus H with parainfluenza virus HN, influenza virus neuraminidase (N), bacterial neuraminidases, eukaryotic neuraminidase and protozoa transneuraminidases, which in itself have very low sequence homologies (Table M). The crystal structure of the neuraminidases of influenzaviruses A and B, salmonella typhimurium LT2 and vibrio cholerae show the same fold and a remarkable similarity in the spatial arrangement of the catalytic residues, although the sequence similarity is low (4, 11, 12, 56). Seven active site residues are common in most of these neuraminidases: R118, D151, E277, R292, R371, Y406 and E425 (1-7, respectively) according to the numbering of influenzavirus A/Tokyo/3/67 (58). All resolved neurami-

nidase structures are organized as a so-called propeller. The 3D models here provided describe a β-propeller which is a superbarrel comprising six similarly folded antiparallel β-sheets of four strands each. In the superbarrel the six sheets are arranged cyclically around an axis through the centre of the molecule like the blades of a propeller. The centre of the molecule forms the active site and binds sialic acid. Also the way the sheets are connected is conserved: the fourth strand of each sheet is connected across the top of the molecule to the first strand of the next sheet (see figure 1). The invention provides the exact position of these sheets, and more importantly, the detailed secondary structure and amino acid sequences of the individual strands and loops composing the sheets of the various viruses. A general notation for the secondary structure elements of the subunit is $\beta_i S_j$ or $\beta_i L_{mn}$ where i = 1 to 6 for the six ß-sheets, j = 1 to 4 for the four strands per sheet and where the loop structures are designated L_{01}, L_{12}, L_{23} and L_{34}, which refer to, respectively, the loop connecting strand four of the preceding sheet with strand one of the next sheet (L_{01}), and the loops connecting strand one with strand 2 (L_{12}), strand 2 with strand 3 (L_{23}), strand 3 with strand 4 (L_{34}). Loops L_{01} and L_{23} all protrude from the top surface and loops L_{12} and L_{34} are all on the bottom surface. The antigenic or immunodominant sites and epitopes composed by these loops all protrude from the top surface. In Tabels A to L, the amino acid sequences of various *paramyxovirinae* are listed, but subsequent alignment of yet unaligned paramyxovirinae sequences or of sequences of yet other neuraminidases is now also within reach of the ordinary skilled researcher wanting to select additional antigenic or immunodominant or catalytically active (neuraminidase like) sequences. The invention thus provides an isolated or recombinant proteinaceous substance comprising at least one virus epitope derived from the attachment protein of a virus from the

family of paramyxoviridae, said epitope corresponding to an antigenic site present on the HN protein of paramyxovirus, which site is identified as one of loop β1L01, β1L23, β2L01, β2L23, β3L01, β3L23, β4L01, β4L23, β5L01, β5L23, β6L01 and β6L23, or a functional equivalent thereof. Such immunodominant or antigenic sites or epitopes can be used alone as (synthetic) peptide, or in combination or in line with other sequences or recombinantly expressed in vaccines specifically directed against paramyxovirinae infection. Such vaccines can be aimed at viruses, such as measles or mumps, causing human disease, but also at viruses causing disease in other animals, such as rinderpest or canine distemper. The invention also provides a synthetic or natural (monoclonal) antibody specifically directed against a virus epitope derived from the attachment protein of a virus from the family of paramyxoviridae, said epitope corresponding to an antigenic site present on the HN protein of paramyxovirus, which site is identified as one of loop β1L01, β1L23, β2L01, β2L23, β3L01, β3L23, β4L01, β4L23, β5L01, β5L23, β6L01 and β6L23, or a functional equivalent thereof. The invention further provides a method for selecting, identifying and producing an epitope of yet another virus of the family of the paramyxoviridae, comprising aligning the sequence of the HN protein of said virus with the sequence of a second virus of the same family of which the 3-D structure is given by the invention, identifying the sequence of said first virus which corresponds with an epitope of said second virus and synthesizing or isolating a proteinaceous substance having said sequence or a functional equivalent thereof. A preferred virus to use as a second virus in such a method is a paramyxovirus such as the bPIV-3 virus. In addition, the invention provides a marker vaccine comprising a virus from the family of paramyxoviridae which virus is modified by functionally removing an immunodominant epitope, which immunodominant epitope corresponds to one of the following

sites which are identified as one of loop β1L01, β1L23, β2L01, β2L23, β3L01, β3L23, β4L01, β4L23, β5L01, β5L23, β6L01 and β6L23. Such a marker vaccine which is altered in one or more immunodominant sites as compared to its wild type virus can be used to vaccinate against a specific disease resulting in vaccinated animals which can easily be differentiated from the wild-type infected animals by having a different immune response. The invention also provides a diagnostic test comprising an immunodominant epitope or loop provided by the invention and/or an antibody specifically directed thereto. The production of such diagnostic tests is witin the skill of the art, using enzymes and/or chromophores known in the art. Diagnostic tests accompanying marker vaccines are provided by the invention since relevant immunodominant sites are now known. The 3-D model of morbillivirus H disclosed by the invention identifies yet unknown neuramidase glycosidic activity for *Morbiliviruses* measles, CDV, RPV, PDV and PPRV. Active site residues in *Morbiliviruses* correspond to active site residues 1, 5, 6, 7 of the aligned neuraminidases: R118, R371, Y406 and E425 according to the numbering of influenzavirus A/Tokyo/3/67 and to R106, R533, Y551, and E569 of measles (figure 5). The invention provides a substance partly or wholly blocking this previously unknown enzymatic activity of the *morbillivirus* H protein. An example of such a substance according to the invention is a carbohydrate, often possibly attached to a (poly)peptide or amino acid, such as a sialic acid, being a possible diacetyl derivative and/or having N-glycolyl groups. An example provided by the invention is a sialic acid modified at the 5 or 6 position. The invention also provides a pharmaceutical composition comprising a substance as identified above mixed with a pharmaceutically acceptable carrier. Such a composition can be used as a prevention or therapeutic medicament for (respiratory) disease, for instance with measles. The invention also provides a method identifying a substance as

above comprising using a morbillivirus, or morbillivirus H protein, e.g. in a neuraminidase test. The invention is explained more in detail in the exemplary part of this description, which, however, should not be seen as limiting the invention.

EXPERIMENTAL PART

Amino acids are abbreviated according to the one letter code.
A:Ala; C:Cys; D:Asp; E:Glu; F:Phe; G:Gly; H:His; I:Ile; K:Lys; L:Leu; M:Met; N:Asn; P:Pro; Q:Gln; R:Arg; S:Ser; T:Thr; V:Val; W:Trp; Y:Tyr.
H: Hemagglutinin
N: Neuraminidase
HN: Hemagglutinin-neuraminidase
RPV: rinderpestvirus
PPRV: peste des petits ruminants virus
PIV: parainfluenzavirus
CDV: Canine distemper virus
PDV: Phocine distemper virus
NDV: Newcastle disease virus
SV: Sendai virus
BRSV: bovine respiratory syncytial virus
MAb: monoclonal antibody
DANA: N-acylneuraminidase inhibitor, 2,3-dehydro-2deoxy-N-acetyl-neuraminic acid.

MATERIALS AND METHODS:

Based on the conservation of neuraminidase active site residues in influenza virus neuraminidase (N) and paramyxovirus hemagglutinin-neuraminidase (HN) it has been suggested that the three dimensional structures (3D) of the globular heads of both proteins are broadly similar. In this study, details of this structural similarity are worked out

and epitopes, antigenic sites and catalytic sites are identified. Multiple sequence alignment of paramyxoviruses HN was used as an intermediate to align the morbillivirus hemagglutinin (H) proteins with neuraminidase. 3-D structures were built for *paramyxovirus* HN and *morbillivirus* H based on homology modelling. Location of insertions and deletions, glycosylation sites, active site residues and disulfide bridges agree with the proposed 3-D structure of the *paramyxovirinae* HN and H. Moreover, details of the modelled H protein show previously undescribed enzymatic activity. This finding was confirmed for Rinderpest virus and Peste des Petits Ruminants virus. The enzymatic activity was highly substrate specific because sialic acid was only released from crude mucins isolated from bovine submaxillary glands. The enzymatic activity indicates a general infection mechanism for respiratory viruses and the active site is a new target for antiviral compounds or substances.

Cells and viruses:

Rinderpest virus (RPV), strain RBOK, peste des petits ruminants virus (PPRV), (kindly provided by Dr. J. Anderson, Pirbright, UK), measles virus (MV), strain Edmonston (16), phocine distemper virus (PDV), strain 1-3, fourth passage, canine distemper virus (CDV), strain rockborn, first passage, and dolphin morbillivirus (DMV), strain 16A, seventh passage. (MV, PDV, CDV and DMV were kindly provided by Dr. A.D.M.E. Osterhaus, Erasmus University Rotterdam, The Netherlands), and bovine respiratory syncytial virus (BRSV), strain RB94, were grown on Vero cells. Infected cell cultures were maintained in Eagle's minimum essential medium (MEM) with 2 % fetal bovine serum. Virions were obtained by clarification of tissue culture medium. Further purification of the virions was performed by pelleting the clarified medium through a 40 % sucrose cushion at 250,000 g for 20 minutes. In some

experiments the clarified medium was pelleted without sucrose at 53,000 g for 2 hrs which gave the same results.

Sequence analysis

Multiple sequence alignments were performed using the Pileup program of the Genetic Computer Group (14), which was obtained from the CAOS CAMM Centre in Nijmegen, the Netherlands. Several scoring matrices were used: Dayhoff matrix based on mutations in protein families, and a structural matrix based on possible dihedral angles a residue can adopt in folded proteins (45). Multiple sequence alignments were performed using several representative sequences of neuraminidase family members and morbillivirus H proteins, because the use of a broad family of homologous sequences improves the accuracy of structure predictions. However, simply aligning by following the rules of the computer programme did not result in a useful comparison; Only after a large number of gaps and alignments had been introduced, the models of the invention were obtained. Secondary structure predictions were performed using the neural network-based program PHD (51), which was obtained from the European Molecular Biology Laboratories in Heidelberg, Germany. The following neuraminidase, HN or H sequences were obtained from the CAOS CAMM Centre for analysis and comparison: *Vibrio cholerae* neuraminidase, strain Ogawa, (accesion number P37060); *Actinomyces viscosus* neuraminidase, strain DSM 43798, (S20590); *Trypanosoma cruzi* flagellum-associated protein, (S32016); *Salmonella typhimurium* neuraminidase, strain LT2, (P29768); *Clostridium septicum* neuraminidase, strain NC 0054714, (P29767); rat cytosolic neuraminidase (42); influenza A neuraminidase, strain A/NT/60/68, (A00885); influenza B neuraminidase, strain B/Beijing/1/87, (B38520); human parainfluenza 2 hemagglutinin-neuraminidase, strain Toshiba, (A33777); Newcastle disease virus hemagglutinin-neuraminidase, strain Beaudette C/45, (A27005); Sendai virus

hemagglutinin-neuraminidase, strain HVJ, (A24004); Bovine parainfluenza 3 hemagglutinin-neuraminidase (B27218), Canine distemper virus hemagglutinin, strain Onderste-poort, (A38480); Measles hemagglutinin, strain Edmonston, (A27006). Molecular modelling was performed using software of SYBYL version 6.0 (Tripos Associates, St. Louis, MO) on a Silicon Graphics Indigo. Energy minimization was performed using the Tripos Sybyl version 6.0 force field. Minimization was performed using a dielectric constant e = 1. Minimization was performed in stages using steepest descent and conjugate gradient; each stage the atoms were given more freedom as described (33). For the introduction of insertions and deletions in the model structure, the program LOOP SEARCH in the SYBYL package was used. The loop regions were taken from a protein fragment database and the selection was based on the correct length, maximum amino acid homology and minimum Root Mean Square difference of the anchor residues in the start and the end of the loop.

Neuraminidase assays.

Neuraminidase assays were performed as described by Aymard-Henry et al. (1) using different morbilliviruses grown on Vero cells. 50 ml Purified virus was added to 50 ml of the substrates and 100 ml buffer for 18 h at 37°C. The following substrates were tested for sialic acid release: fetuin from fetal calf serum (M2379, Sigma, St.Louis) at 50 mg/ml; mucin type 1, isolated from bovine submaxillary glands (M-4503, Sigma, St. Louis) at 50 mg/ml, (in some experiments 100 mg/ml was used, followed by a clarification of the solute); mucin type 1-S, isolated from bovine submaxillary glands and further purified (M-3895, Sigma, St.Louis) at 50 mg/ml; mucin type 2, isolated from porcine stomach (M-2378, Sigma, St. Louis) at 50 mg/ml; 6'-N-acetylneuramin-lactose from bovine colostrum (A-8556, Sigma) at 10 mg/ml; 3'-N-acetylneuramin-lactose from bovine colostrum (A-8556, Sigma) at 10 mg/ml;

bovine hyaluronic acid (H-7630, Sigma) at 50 mg/ml; human hyaluronic acid (H-1751, Sigma) at 50 mg/ml. Neuraminidase from *Clostridium perfringens* (N-2876, Sigma) was used as a positive control.

Haemadsorption test

Haemadsorption tests were performed as described at 20°C (57). Vero cells were infected with measles virus strain Edmonston. Erythrocytes of Cercopithecus monkeys were a kind gift of P. de Vries of RIVM, Bilthoven.

RESULTS:

Multiple sequence alignments (figure 5) were performed with several representative neuraminidase sequences of influenzaviruses A and B and several representative HN sequences of paramyxoviruses in order to combine the two separate multiple sequence alignments of influenza N and paramyxovirus HN described by Colman et al. (9). Alignments were performed using diverse parameters for gap-weight and gap-length-weight and two different similarity matrices (materials and methods). Parts of the computer generated alignments were combined manually in a final alignment in such a way, that the active site residues as described by Colman et al. (9) were now properly aligned. Manual editing of the alignment was also assisted by a neural network-based secondary structure prediction (51) which was occasionally used as a guideline to align sequence "blocks" with low homology. If possible, gaps were avoided in regions corresponding to strands in the influenzavirus N molecule. The alignment between influenzavirus N and parainfluenza HN could not be generated using only the computer but had to be performed visually, because none of the cysteine residues matches between influenza and parainfluenzaviruses. Because these residues have a very high score in the similarity-

matrix, the computer-generated alignment was incorrectly biased, leading to an incorrect 3-D model.

The alignment was extended with the multiple sequence alignment of bacterial protozoa and eukaryotic neuraminidases. The correct alignment between the bacterial and protozoa with the viral neuraminidases was based on the structural alignment of influenza N and *Salmonella typhimurium* N according to Crennel et al. (11), which was based on topologically equivalent residues. Finally, the multiple sequence alignment of paramyxovirus HN was used as an intermediate set of sequences to align the morbillivirus H proteins with all other neuraminidases. ß6S4 of morbillivirus H and paramyxovirus HN are homologous according to a circular alignment. Although the similarity in primary sequence of bacterial neuraminidases compared to the primary sequence of influenzavirus neuraminidase was very low (7.3% - 11%) (Table M), the crystal structure of *V. cholerae* neuraminidase and *Salmonella typhymurium* neuraminidase shows after modelling the same fold as influenzavirus neuraminidase. Correspondingly, the primary sequences of paramyxovirus HN and influenza virus N also have a low similarity (7.1% - 10%) (Table M). The similarity between influenza N and morbillivirus H is even lower (6.3 %) but the similarity between paramyxovirus HN with morbillivirus H is higher (11.4% - 14.5%) than the similarity of paramyxovirus HN with influenzavirus N (7.1 % - 10 %)(Table M). Therefore, the parainfluenza sequences was now used as an intermediate to align the morbillivirus H with the influenzavirus N sequences. Because the structural model of parainfluenza HN is used as an intermediate to build the model of measles - virus H, the largest uncertainty of the model is the similarity between influenza N and parainfluenza HN. The first part of the alignment of morbillivirus H with other neuraminidases or transneuraminidases is complex. Especially the alignment of the first sheet and the location of the

second stem region. According to alignment procedures described in *Materials and Methods*, a global homology is found approximately C-terminally from position 226. However, the highest, most significant local homology was found for $L^{105}R^{106}T^{107}P^{108}$, which is homologous to the most conserved region in all neuraminidases and trans-neuraminidases. To incorporate this best local homology with great functional importance with the best global alignment, an excessive gap had to be introduced in the morbilli virus H sequence alignment. As a result a major part corresponding to the possible parainfluenza stem-region is 'deleted' in the alignment, and a large part is 'inserted' in β1L12. A topology of morbillivirus H is as follows: after the transmembrane region, the first smaller stem insert extends up to the neuraminidase head, the large second insert appears in a loop of the neuraminidase β-propeller, which suggests that right after the first β-strand of the β-propeller, which contains the very important first catalytic Arginine, the polypeptide folds back under the neuraminidase head to form a stem together with the smaller insert, and than the chain returns to continue the β-propeller (Fig. 1). The relatively large deletions in sheets β5 and β6 typical for morbillivirus H, may be a consequence of the bulky stem region of morbillivirus H and perhaps Cys606 is connected to a cysteine in the stem. Using the alignment, a three-dimensional (3D) model was constructed of a paramyxovirus HN by replacing the residues of the crystal structure of influenza N with the homologous residues of bovine parainfluenza-III (bPIV-3). For residues contained in gap regions according to the alignment, loop searches were performed as described in materials and methods. The large and small loops that were constructed in this way were not foreseeable by computer analysis. The loops were chosen from a group of loops which were selected on

basis of homology and distance of the anchor residues in the start and the end of the loop. The final choice is based on structural limitations in the three-dimensional space of neighbouring loops, positioning of important residues in the active site or close spatial positioning of cysteine residues which likely form a disulfide bridge. Finally the structure was minimized. Similarly, a 3-D model of measles H was built according to the alignment of bPIV-3 HN and measles H as described above. The modelled 3D structure of bPIV-3 HN was used as a framework for the homology modelling of measles H.

Location of insertions and deletions in paraflumorbillivirinae HN and H.

The reliability of the model is strengthened when insertions and deletions occur at appropriate locations. According to the alignment and the model, the large insertions (β1L01, β2L01, β2L23, β3L01, β5L01 and β5L12) and the very large insertions (β3L23, β4L01) are all located on the top of the barrel, except for β5L12. This is in accordance with the general neuraminidase-fold in which the top-loops (L01, L23) are always extensive compared with the bottom loops (L12, L34). In contrast, the large deletions: (β1L23, β5L01, β5L34, β6L12) seem equally distributed over the top or bottom side of the barrel. The bottom side deletions are found in the C-terminal part of the barrel in sheet β5 and β6 and are larger for morbilliviruses.

The alignment is constructed from four groups of different multiple sequence alignments which display a low homology between the groups (Table M). Reliable alignments display a high homology and few gaps. Because the number and length of gaps will affect the quality of an alignment, the introduction of gaps in an alignment is not favourable. However, the similar location of some gaps in independently aligned groups of the total alignment, reinforces the quality of the

alignment. Thus, some of the insertions introduced in the alignment of influenzavirus N with *paramyxovirinae* HN/H are more acceptable because they appear in regions which also show gaps in another group of the total alignment. For example an insertion is present in ß1L01 of Paramyxovirus HN, and a much larger insertion is present in β1L01 of bacterial/protozoan N. Similar equal insertions or deletions present in paraflumorbilli HN/H and in bacterial/protozoan N occur in β1L23, β2L01, β2L23, β3L01 and β4L01.

Disulfide bridges

According to the 3D structure of bPIV-3 HN and measlesvirus H, the cysteine bridge pairing can be obtained. Strikingly, there is no single conserved cysteine bridge between influenzavirus N and *paramyxovirinae* HN or H. One exception may be a cystine bridge between βS2 and β6S3 in influenzavirus N and morbillivirus H, but even this bridge is not structurally similar because in N the start of S2 is connected to the end of S3 although in H, the end of S2 is connected to the start of S3. All cystine bridges in the morbillivirus model are conserved compared to the cystine bridges in parainfluenza HN, except for the cystine bridge between ß6S2 and ß6S3 in morbilli virus H.

Cystine bridges between residues 159-571, 190-214, 204-265, and 535-544 in parainfluenza virus HN were already predicted by Colman et al. (9).

A study on the role of the individual cysteine residues in the HN protein of NDV (40) suggested that, (according to bPIV-3 numbering), cysteines 190 and 214 are linked; cysteines 204, 256, 265 and 269 are linked in some way; cysteines 363, 463, 469 and 473 are linked in some way and cysteines 535 and 544 are linked. These results agree with our model if bPIV-3 and NDV show identical cysteine bridge connections.

Tetramer interface

The HN and H proteins are thought to form tetramers as mature proteins (7, 37, 41, 44, 55). A model of the tetramer was generated by superimposing the monomer models on the backbone of the influenza neuraminidase tetramer. The two largest insertions (> 15 residues) are located on β3L23 and β4L01 which agrees with the tetramer model, because these loops are on the outside of the tetramer, away from the interfaces. The only region that seems to obstruct an appropriate tetramer formation is the inserted β2L01 loop. Therefore, in the actual structure, β2L01 must be located more towards the active site. Most conserved non-charged residues in measles H are located on β-sheets 1 and 2 which form part of the tetramer interface.

Glycosylation sites

The potential glycosylation sites in the model of bPIV-3 HN are located on the surface and mostly on loops on the top of the molecule. For bPIV-3 HN, the potential glycosylation sites are located on β3L01, β3L23, β5L01, β6L01 and on β6S3 and β6S4. β6L01, β6S3 and β6S4 cluster in the 3-D space. However, β6S3 is less likely to be used because a carbohydrate at this site may obstruct tetramer formation. Two potential paramyxovirus glycosylation sites that have no direct counterpart in bPIV-3 reside on β4L01 (mumps and PIV-2) and β5L23 (sendai, PIV-2 and mumps) (Fig. 5a in yellow). The first is very close to the potential glycosylation site on β5L01 of bPIV-3 and the second is very close to the potential glycosylation site on β6L01 of bPIV-3. β3L01 is very close to a N-linked carbohydrate in the structure of influenza A neuraminidase on β2L23. Strikingly, most potential glycosylation sites are located away from the tetramer interface. For NDV HN the actual usage of sites has

been determined (39). Sites 2 (β3L23), site 3 (β4S4) and site 4 (β5S2) can be accomodated in the structure and are located at the side or bottom of the tetramer, away from the teramer interface. Site 5 (β5L23) is hidden in the interior and is located right under an active site residue. A glycosylated site 6 (β6L12) might interfere with tetramer formation. This might explain why potential glycosylation sites 5 and 6 are not used (39).

For measles, most potential glycosylation sites are located on the postulated stem-region. Only one potential glycosylation site, which is not used in this strain (21), is located on the neuraminidase head on loop β1L23 (Fig. 5b in purple). The corresponding loop in influenza A and B neuraminidase contains the only conserved glycosylation site. Three potential morbilli virus glycosylation sites on H that have no counterpart in measles reside on β3L23 (RPV and PDV), β4L01 (PDV and CDV) and β6S4 (PDV and CDV), all of which have counterparts in paramyxovirus HN.

Epitopes

The bPIV-3 HN model can be used as a general model for paramyxovirus HN. Therefore, antigenic sites of all HN proteins can be used for localizing the epitopes on the 3D model of bPIV-3 HN. Indeed, several individual loops correspond to previously identified immunodominant regions in individual viruses, demonstrating the strength of the overall models provided by the invention. Loop ß1L23 corresponds to antigenic site 23 in NDV HN as described by Iorio et al. (24). Antibodies against antigenic site 23 recognize only the oligomer (38) which agrees with the location of β1L23 which is close to the tetramer interface, which is in agreement with competition studies (25).

Antibody escape mutants with substitutions at residue positions 363 and 472 of SV-5 were selected by antibodies

directed against antigenic site 4 (2). According to the alignment, the mutations are located on β3L23 and β5L01 next to a postulated disulfide-bridge, corresponding to the bPIV-3 HN model. The vicinity of both mutations within the discontinuous epitope agrees with the tertiary structure of the HN model. Substitutions disrupting the binding of antibody directed against antigenic site 5 of SV-5 HN occur at positions 453, 498 and 541. In the model these residue positions are located on β5L01, β5L23 and β6L12 respectively. Residues 453 and 498 of SV-5 are located on top of the molecule on two neighbouring loops according to the bPIV-3 model (Ca atoms within 11 A). However, residue 541 of SV-5 is located on the bottom side of the molecule. Otherwise, antigenic site 5 may map to the side of the molecule and span from the top to the bottom of the molecule. Alternatively, a mutation which structurally compensates for a harmful mutation may itself lie outside the antigenic site recognized by the selecting mAb.

In hPIV-3 HN, mutation of residue 281 or 370, and 278 disrupt binding of antibodies against overlapping epitopes I and VI, respectively (6). Residues 278 and 281 are located on the exposed surface loop β2L23. However, residue Proline 370 is located 30 Å away on β3S3, and is not exposed. Perhaps mutation of residue 370 can allosterically induce a conformational effect on the epitope.

Because the antigenic regions of measles virus H protein have been studied extensively, these data are very useful to check the validity of the measles H model provided by the invention. Comparison of the location in the 3D model of epitopes mapped with the aid of short synthetic peptides (35, 36), showed that already known antigenic sites are located on β1L23, β2L01, β2L23, β3L23 on the top of the molecule and β4S2L23 and β6L34S4. All of these regions except for β4S2L23 are in agreement with the model because they are exposed on the protein surface.

With the aid of five mAbs, four antigenic sites could be characterized on measles H protein (21, 53). For these mAbs, (Mab I-29 to site I, Mab 16-CD11 to site II, Mabs 16-DE6 and I-41 to site III, and Mab I-44 to site IV), the epitopes were determined by sequencing selected monoclonal antibody resistant mutants. Mab I-29 maps to residues 313 and 314 on a large insertion on top of β2L23. This epitope was also mapped with peptide binding studies (35, 36). Mab I-41 maps to residue F552, the first residue on strand β6S1 in the centre of the molecule right under active site residue Y551. Mab 16-CD11 maps to residue 491 at the centre of the large loop ß5L01. Mab 16-DE6 mapped to residues 211, 388, 532 and 533 (21). Because residue 211 lies outside the model somewhere between β1S1 and β1S2, the spatial relationship with other residues in the same antigenic site can not be verified with the 3D model. Residues 388 and 532, 533 are located on top of the molecule on loops β3L23 and β5L23 respectively, and therefore this discontinuous antigenic site supports the model. According to Liebert et al. (33) the major antigenic site of measles H protein is located between residues 368 and 396, which corresponds exactly to the large insertion at β3L23.

Active site paramyxovirus:

The alignment predicts that six of the seven common active site residues are conserved in paramyxovirus HN. The active site residue influenza-D151 has no homologue in paramyxovirus HN according to the alignment. Residue influenza-D151 is probably involved in proton transfer, however the enzyme is active above the pK_a of D151. So a non-specific proton donor, like a water molecule, may be involved (4). Influenza-D151 aligns with parainfluenza-Q222, but Q can not act as the proton donor. As mentioned above, the role of influenza-D151 is still obscure, the conservation in influenza and some bacterial neuraminidases suggests an important function, but

according to the sequence alignment the aspartic acid is also not conserved in *Streptomyces lividans*, (M. viridifaciens) and *Actinomyces viscosus*. If an aspartic acid is the proton donor, than two candidate residues can be conceived: parainfluenza-D216 in ß1L23 or parainfluenza-D279 in ß2L23. In the case of parainfluenza-D216, the alignment needs minor justification, in the case of parainfluenza-D279, loop ß2L23 has to be remodelled for the correct orientation of D279 in the active site.

The most conserved region of paramyxovirus HN corresponds to the ^{252}NRKSCS^{257} sequence, located on ß1L01-ß2S1. The region corresponds to the only sheet in influenza that does not contain active site residues. Parainfluenza-R253 which is part of the highly conserved stretch NRKSCS may be homologous to the conserved influenza-R152. In that case, parainfluenza-R253 is not homologous to influenza-R224 as suggested by Colman et al. (9), but instead the positively charged parainfluenza-K254 may be homologous to influenza-R224. Influenza-R152 has an important active-site structural role because it directly contacts N-acetyl of sialic acid, while influenza-R224 is just a framework residue which holds influenza-E276 in place. Because there is no homologue for influenza-E276 in HN, such a framework function is not expected in parainfluenza HN. Perhaps parainfluenza-K254 holds active-site residue parainfluenza-E409 in place. Parainfluenza-R411 and D480 are conserved charged residues, close to active site residue Y530, without counterparts in influenza N. Perhaps, parainfluenza-R411 is a framework residue for active site residue parainfluenza-E409 or it may contact parainfluenza-D480. As suggested by Colman et al. (9), parainfluenza-D480 may be a framework residue that binds parainfluenza-R424 (fourth active site residue).

Morbillivirus:

After close inspection of the location of all conserved charged residues in the 3D model of measles H, we noticed that most conserved charged residues are clustered at the top-centre of the β-propeller, where the active site is located in neuraminidases. The clustering of conserved charged residues is suggestive for a conserved glycosidase activity in measles H. The highest conservation was observed for amino acids that are in close proximity to the glycoside bond of sialic acid. Although for H of morbilli viruses only hemagglutination and no neuraminidase activity has been reported, the finding (provided by the invention) of some conserved active site residues surprisingly shows that H has enzymatic activity.

The alignment predicts that four of the seven common active site residues, as described in the introduction, are conserved in morbillivirus H. measles-R106 is homologous to influenza-R118, measles-R533 is homologous to influenza-R371, measles-Y551 is homologous to influenza-Y406 and measles-E569 is homologous to influenza-E425. The conservation of both measles-R106 and measles-E569 is coherent because these two residues form a conserved 'couple' in neuraminidases important for the catalytic mechanism (4). According to the alignment, the conserved residue measles-R533 has a very important role in substrate binding, it binds the acidic group of sialic acid and is responsible for the precise orientation of the sugar for the glycosidic cleavage. Measles-Y551 is one of the most important residues in the reaction mechanism because it stabilizes the oxocarbonium intermediate. Furthermore, two additional active site residues are conserved: measles-R253 is homologous to influenza-R152 and measles-N450 is homologous to influenza-N294.

In general, no homologies are observed for the side of the active site that interacts with the sialic acid glycerol sidechain in influenza neuraminidase.

According to the alignment there are no homologues in morbillivirus H for the typical active site residues 2, 3 and 4 corresponding to influenza-D151, -E277, and -R291, respectively. According to the alignment, active site residue 2 is also not present in the neuraminidases of paramyxovirus, streptomyces lividans, and actinomyces viscosus. The active site residues 3 and 4 are also not present in the neuraminidase of Trypanosoma cruzi listed in Table M. In contrast to the alignment between paramyxovirus HN and influenzavirus N, the missing aspartic acid of active site residue 2 can not be solved by a justification of the alignment. The third active site residue of influenza-E277 on β4S1, which is missing in morbillivirus H, has an important role in the neuraminidase mechanism of influenza because it accepts a proton from active site residue influenza-Y405. In the 3D space, this active site residue may be substituted by another proton acceptor. The negatively charged conserved residues D505 and D507 are located on an insertion on loop β5L01. The important location and conservation indicate a role for these residues in the active site. There are no obvious homologues for these residues in other neuraminidases according to the alignment, thus D505 or D507 substitute for the missing active site residue corresponding to influenza-E277 and the framework residue corresponding to parainfluenza-D480.

One of the few candidates for the missing fourth active site residue (influenza-R292 on β4S2) is measles-R547 on β6L01. However, the 3D model does not support a superposition of influenza-R292 and measles-R547.

A remarkable conserved cluster of residues in morbillivirus H are Q109 on β1S1 and H354, R355 on a characteristic β-bulge on β3S1. The residues are in close proximity to the ligand binding site. The residues approximate the 3D space occupied

by conserved negatively charged residues in β2S1 or β3S1 in bacterial and influenza neuraminidases respectively. The role of the residues is unknown but their location and conservation suggest a possible role in proton transfer. Several conserved negative charged residues are found near the ligand binding site of measles: measles-E256, measles-D530 and measles-D574 on respectively β1L23, β5L23 and β6L23 but non of these are superimposable on influenza-D151.
G432, P433, I435 are conserved non-charged residues on β4S1 at the bottom of the active site. These residues are very close to P368 on β3S2 which also lines the active site pocket. On the other side of the active site, the conserved G104, L105, P108 and Q109 on β1L01 and β1S1 line the pocket.

Neuraminidase assays

Because the 3D model of morbillivirus H suggested a neuraminidase activity that has never been described before, neuraminidase assays were performed with rinderpest virus (RPV) and a large selection of neuraminidase substrates (Fig. 2). Sialic acid was only released from mucin type 1, isolated from bovine submaxillary glands. Figure 3 shows that the neuraminidase activity of RPV was dose-dependent and no activity was found in supernatants of mock-infected or BRSV-infected cells. Next, MV, PDV, CDV, DMV and PPRV were tested for neuraminidase activity. Only PPRV showed a low neuraminidase activity, only with bovine submaxillary mucin type 1 (Fig. 3). Neuraminidase from *Clostridium perfringens* showed good activity with mucin type 1 (data not shown) but bovine PIV3 did not. Neuraminidase activity of RPV could be inhibited to 2.9% by preincubation of RPV with an RPV-specific polyclonal cow serum.
The optimal pH for the RPV-associated neuraminidase is shown in Fig. 4. The activity of RPV neuraminidase extends a relatively wide and acidic pH-range with an optimum between pH 4 and pH 5, which is typical for viral neuraminidases.

Neuraminidase activity was reduced to 50 % after incubation of the virus at 61°C for 25 minutes, and neuraminidase was completely inactivated after heating at 100°C for two minutes.

Activity could not be inhibited by the N-acylneuraminidase inhibitor, 2,3-dehydro-2deoxy-N-acetyl-neuraminic acid. Like paramyxovirus HN, the neuraminidase activity of RPV was independent of Calcium (data not shown).

Haemadsorption assay

At 20°C Cercopithecus erythrocytes adsorbed in a single cell circle around cpe spots of a measles infected monolayer. After incubation at 37°C for 3h, in order to activate the presumed neuraminidase activity, all adsorbed erythrocytes were detached from the monolayer. Subsequently, the infected cells were still able to adsorb erythrocytes at 20°C.

Discussion

In this study, 3D structures are identified for the attachment proteins of all *paramyxovirinae*. The structure of influenza N was used as a framework for modelling the paramyxovirus HN, extending a previous proposal for the gross structural arrangement of this protein (9, 26). The neuraminidase multiple sequence alignment could be extended with morbillivirus H sequences when the paramyxovirus HN sequences were used as an intermediate. Consequently, a 3D structure of morbillivirus H could be modelled on the parainfluenza HN framework.

Most insertions in the larger neuraminidase heads of paramyxoviruses are located in loops at the top surface of the molecule. This is in accordance with the general neuraminidase-fold in which the top-loops are always extensive and more variable compared with the bottom loops. Additionally, most published experimental data agreed with

the 3D model. Most epitopes are located on the top-loops of the paramyxovirus heads. Multiple mutations in discontinuous epitopes, which were scattered over a large part of the primary sequence, were close in the 3D model. The large insertions are located at sites that are not in the interface of the possible tetramerization sites of the molecule. The potential glycosylation sites were mostly located at the molecular surface and some sites had counterparts in the influenza neuraminidase molecule. Although the cysteine residues were not conserved in the alignment, all residues could be paired in cystine bridges which is a strong support for the 3D model. Finally, the spatial arrangement of proposed active site residues are similar as the active site residues seen in other neuraminidases of known structure. Moreover, the model implies which additional, hitherto unrecognized residues are important for neuraminidase activity or active site structure. For measles (morbillivirus) H, this shows hitherto undescribed enzymatic activity. The prediction embarked a search for the right substrate to prove glycosidic activity in a morbillivirus. Eventually, neuraminidase activity was found for RPV and PPRV with mucin isolated from bovine submaxillary glands. Furthermore, the temperature-dependent haemadsorption of measles further shows its neuraminidase activity.

The alignment of paramyxovirus HN differs from the approximate alignment of Colman et al. (9). They proposed an insertion in HN corresponding to sheet 3 of influenza N, and the start of sheet 3 differs by about half the length of the sheet compared to the alignment described in this study. Because the higher homology between *paramyxovirinae* HN and H sequences, alignment of HN and H was easier compared with the alignment of these proteins with other neuraminidases.

The models illustrate the diverse solutions for the elevation of a neuraminidase head above the viral membrane. In the case of influenza, the neuraminidase 'head' is extended above the

membrane by a stalk region of approximately 40 amino acids. The stalk lifts the neuraminidase head approximately to the same height as the other viral membrane protein: the hemagglutinin, which contains membrane fusion activity. There is no indication for a stalk region in *paramyxovirinae*. However, the corresponding region in paramyxovirus HN, between the transmembrane region and the neuraminidase head, contains a large protein domain (between residue 56 and 161) which has high alpha-helix propensity according to neural network-based secondary structure predictions (data not shown). It is likely that this is a helical stem region that supports the neuraminidase head and lifts it to the same height as the fusion protein, comparable to influenza. According to the unusual alignment, morbilliviruses have acquired a completely different helix-rich domain which is made up of two insertions (a 40 residue insert between residue 58 and 98, and a 110 residue insert between residue 115 and 225, of which the large domain is inserted inside the neuraminidase head domain instead of N-terminal to the neuraminidase head as observed in influenza and paramyxovirus (Fig. 1). Although the "two inserts scenario" in morbillivirus H is not elegant, it is the only way to combine the highest local and the highest global similarities in the alignment. Analogous to the stem region of paramyxovirus HN, also both stem region insertions of morbillivirus H have high helix propensity according to neural network-based secondary structure predictions. Excessive insertion within a neuraminidase gene is not unique. Within the Vibrio cholerae neuraminidase an insertion of a lectin-domain of 193 residues has occurred (10) in ß3L01, between sheet 2 and 3.

Apart from the stem region, the most excessive insertion in *paramyxovirinae* HN/H is the 28 - 36 residues long insertion in ß3L23. This region is the most immunodominant region of measles H. The presence of this large insertion in all *paramyxovirinae* HN or H proteins compared with influenza or

bacterial/protozoan N proteins, the lack of any active site residues in the loop and the antigenicity of the loop supports a possible role as a surface exposed receptor binding site for β3L23. Interestingly, for measles this region is a neutralization site (59) and may play a role in neurovirulence of the virus (33). Measles H and some paramyxovirus HN proteins contain a cystine noose in B3L23 between Cys381 and Cys386. Such nooses are often involved in protein-protein interaction (30). According to Ziegler et al. (59) a cystine noose is present between Cys386 and Cys394. In that case, the structurally ill-defined loop should be remodelled to allow a 386-394, and a 381-494 pairing. Functional studies with chimeric measles H protein showed that residues 491, 493, 495, 505 and 506 may be involved in agglutination to red blood cells because these mutations abrogated binding of a mAb directed to a non-agglutinating H protein (23). Shibahara et al. (54) showed that residue 546 is involved in agglutination. These studies suggest that the two adjacent loops β5L01 and β6L01 are involved in binding activity of measles H. It is presumed that measles virus H binds erythrocytes through CD46 (15, 43). However, freshly isolated wild-type strains do not interact with CD46 (32). This suggests that wildtype measles virus uses a different receptor to initiate infection (32). Use of multiple receptors has previously been described for HIV-1 (17) and may likely be general in virus infections. At least three ligands for H may play a role during infection: sialic acid via the center of the β-propeller during transport through the mucus layer; CD46 via β5L01 and β6L01 for attachment to cells and additionally a possible interaction with F via an unknown site.

The neuraminidase gene is probably spread from eukaryotic cells by horizontal gene transfer among bacteria, fungi and protozoa during association with their animal hosts (50). It is not known whether viral neuraminidase genes also have an

eukaryotic origin. A recently cloned eukaryotic neuraminidase gene for rat cytosolic neuraminidase has a very weak homology with bacterial and protozoan neuraminidases (42). Maybe new eukaryotic sequences will bridge the distances between the neuraminidase superfamily members. In contrast to bacterial and protozoan neuraminidases, viral neuraminidases are transmembrane proteins and they are organized as tetramers. The viral proteins do not possess the "Asp-box" motif (Ser/Thr-Xaa-Asp-[Xaa]-Gly-Xaa-Thr-Trp/Phe), and especially the influenza and paramyxovirus neuraminidases contain more cystine bridges than bacterial and protozoan neuraminidases. Perhaps the viral neuraminidases are examples of unique convergent evolution but if the neuraminidase gene is transferred form a higher organism to the virus, than several evolutionary scenarios are possible for an archetypal myxovirus. It is possible that the archevirus may have possessed an attachment protein that was lost, or changed radically, after the introduction of the neuraminidase gene. Otherwise, the archevirus possessed just one membrane protein: the fusion protein. For its proper function, the introduced neuraminidase acquired several characteristics as mentioned above: a transmembrane region and a tetrameric organization, cystine bridges and an extension of the neuraminidase head to lift it to the same height as the other membrane protein with which it evolved a probable cooperation as is shown for some *paramyxovirinae* (3, 5, 13, 20, 52). Based on the general low amino acid homology between influenza N and *paramyxovirinae* HN and H, and especially the divergence in cystine bridge connection, it is likely that influenza and paraflumorbilli neuraminidases are not evolutionary related. Additionally, because the gene is not present in viruses which are evolutionary more related to influenza virus or *paramyxovirinae*, the neuraminidase gene in orthomyxovirus influenza and *paramyxovirinae* may have been introduced independently. Morbillivirus H contains very few

cystine bridges, but most of these cystine bridges are conserved with paramyxovirus HN. Therefore, neuraminidase may have been introduced before the paramyxo-morbillivirus diversification. Thus, it is possible that the neuraminidase gene was introduced in influenzavirus before the diversification of type A and B and the gene was introduced in *paramyxovirinae* before the diversification of respiro, rubula and morbilli viruses. Influenzavirus N, paramyxovirus HN, and morbillivirus H, have independently acquired a domain that elevated the neuraminidase head above the viral membrane. The very dissimilar "stem regions" of paramyxovirus HN compared with morbillivirus H, suggests that the evolution of the stem occurred independently, after shared features like cystine-bridges and the large B3L23 loop had evolved.

We discovered neuraminidase activity in RPV and PPRV. RPV has been suggested to be the archetype morbillivirus (46). This neuraminidase activity is independent of divalent cations, has a pH optimum typical for viral neuraminidases, is not blocked by the most common neuraminidase inhibitor (DANA), and is highly substrate specific. The high substrate specificity may be related with the inability to inhibit with DANA. We detected the substrate in crude mucins from bovine submaxillary glands. The exact type of sialic acid serving as the RPV-H substrate can now be elucidated. Except for a slight activity with PPRV, the other morbilliviruses did not show any neuraminidase activity with this substrate. Perhaps species specific substrates exist for the other morbillivirus neuraminidases, which substrates can now also be elucidated. Most differences in the morbillivirus active sites are found on the opposite face of the glycosidic bond where interactions occur with the sialic acid glycerol side chain in influenza neuraminidase. This suggests that the substrate for morbillivirus neuraminidases may be sialic acids with typical

modifications at the 5 or 6 positions, but also other carbohydrates are possible.

Influenza neuraminidase, paramyxovirus hemagglutinin-neuraminidase, and morbillivirus hemagglutinin are named after their identified properties. However, because we identified neuraminidase activity in a morbillivirus and because hemagglutinin activity has been observed in several influenza neuraminidase proteins (18, 31), the different names for these topological similar proteins is confusing. The general function for all these proteins may be similar and versatile: neuraminidase activity; carbohydrate binding and/or receptor binding; and in some cases binding to the neighbouring fusion protein. The similarities in structure and function may generalize aspects of the infection mechanism. It is likely that microorganisms that infect the respiratory tract must have evolved a way to migrate through the mucus layer. Binding of virus to the receptor is considered to be a multistep process (19). Multiple receptors could be coreceptors and act together, or the receptors may act sequentially. Virus binding might involve a rapid low-affinity interaction to an abundant receptor like terminal sialic acids on mucin polymers. A virus with a hemagglutinin and glycosidase activity like the orthomyxoviruses and *paramyxovirinae* could then roll over or swim through a mucus layer by continuously binding and cleaving sialic acid. At this stage, application of a therepeutic, specific nauraminidase blocking substrate as a medicament will prevent further infection and thus prevent, mitigate or alleviate disease. After browsing the abundant low-affinity receptor environment it might reach the cell surface and find the second high affinity receptor. Sialic acids ensure the viscoelastic properties of mucins. Therefore, mucuous gels may be disintegrated as was shown by the action of purified *S.pneumoniae* neuraminidase on mucins which resulted in a significant reduction in the native viscoelastic properties of the

mucins (48). Perhaps microorganisms that have to cope with the mucus barrier, evolved ways to overcome this barrier with glycosidic activity, esterase activity or mucin-like regions (29). Using this hypothesis, it would be unusual if the morbilliviruses had no glycosidic activity. Like neuraminidase specificity in bacteria, the neuraminidase specificity may be related to the site of infection (10). The 3D models for *paramyxovirinae* HN and H may guide functional studies. The discovered enzymatic activity of some morbillivirus H proteins provides new potential targets for therapeutic drugs.

Figure legends

Figure 1.
Cartoon of global structures of paramyxovirus HN (a), and morbillivirus H (b). Top section indicates the ß-propeller in which the six sheets are shown as rectangles. Stem regions and transmembrane (TM) regions are indicated. Arrows indicate the direction of the polypeptide chain. In (c) the separate loops and strands composing one sheet are shown.

Figure 2.
Neuraminidase activity in RPV was determined using different substrates. RPV was sedimented by ultracentrifugation. 4.6 $TCID_{50}$ RPV in 50 ml was incubated overnight at 37°C with the substrates as in "Materials and Methods".

Figure 3.
Doses-response curves of different morbillivirus dilutions, bovine PIV3 and BRSV.
Concentration of virus (x-axis) is plotted against the amount of released sialic acid from mucin-1 (y-axis).

Figure 4.
Neuraminidase pH optima. Neuraminidase activity of 4.5 $TCID_{50}$ RPV was assayed under various pH conditions.

Figure 5.
3-D models based on alignment of bacterial, protozoan, eukaryotic and viral neuraminidases. In combination with figure 1, the 3-D models are given. The strands of ß-sheet 1 to 6 are colored from purple (N-terminus) to cyan (C-terminus). Strands are colored purple in sheet 1, magenta in sheet 2, red in sheet 3, orange in sheet 4, green in sheet 5 and cyan in sheet 6. In between the strands the loops are indicated in black (for identification see tables A-L). ß-

Strands are assigned according to 3D structures of *S. Typhimurium* neuraminidase and influenzavirus A neuraminidase. Cystine bridges in influenza A and B are shown as green lines. Proposed cystine bridges in *paramyxovirinae* are shown as blue lines. Connections of long range disulfide bonds are indicated with cirlce and box symbol respectively. Proposed disulfide bonds in *paramyxovirinae* are based on structural models in Fig. 3 a,b. Active site residues are shown in red and are numbered 1 to 7. Capital letters are used for conserved residues in respective columns. Unaligned protein domains are indicated as square boxes. TM indicates the transmembrane region. Large inserts are indicated with diamant boxes. Length of inserts are given in parantheses in the boxes. Gaps are indicated by periods. Residue numbering is indicated after each line. Crystal structures are known for neuraminidases of *V. cholera*, *S. typhimurium*, and Influenza A and B.

Table A: Loop β1L01

bpi-iii	PKIRLIPGPGLLATSTTVNGC.IR
sendai	PEISLLPGPSLLSGSTTISGC.VR
hpi-ii	P...LLNIPSFIPSATSPNGC.TR
sim-v-5	P...LLNMPSFIPTATTPEGC.TR
mumps	H...PLNMPSFIPTATSPNGC.TR
ndv	P.SAFQEHLNFIPAPTTGSGC.TR
rpv	.....IGDEVG.LR
measles	.....IGDEVG.LR
pdv	.....IGDEVG.LR
cdv	.....IGDEIG.LR

Table B: Loop β1L23

bpi-iii	GCQD.....IGKSYQVLQI
sendai	GCAD.....IGKSYQVLQL
hpi-ii	DCLD.....FTTSNQYLAM
sim-v-5	GCQD.....HVSSNQFVSM
mumps	NCKD.....HTSSNQYVSM
ndv	GCRD.....HSHSHQYLAL
rpv	NQRARRPSIVWQQDYRVFEV
measles	NLSSKRSELSQLSMYRVFEV
pdv	DTEEN....FETPEIRVFEI
cdv	DIERE....FDTREIRVFEI

Table C: Loop β2L01

bpi-iii	HTFNIDDNRKS
sendai	HTYDINDNRKS
hpi-ii	IYLSDGINRKS
sim-v-5	LYLSDGVNRKS
mumps	QYLSDGLNRKS
ndv	ISLDDTQNRKS
rpv	LELPRQPELET
measles	LEQPVSNDLSN
pdv	RIISNNSNTKI
cdv	MVLPKNSKAKV

Table D: Loop β2L23

bpi-iii	DERSDYASTGIED
sendai	DERTDYSSDGIED
hpi-ii	SEKEDYATTDLAE
sim-v-5	PERDDYFSAAPPE
mumps	LETDDYAGSSPPT
ndv	TEEEDYNSAVPTL
rpv	VALHYGRVGDDNK
measles	ITIPYQGSGKGVS
pdv	ILLNLGDEESQNS
cdv	VLLYHDSSGSQDG

Table E: Loop β3L01

bpi-iii	...NNITFD
sendai	...SEVDLD
hpi-ii	....LPNTT
sim-v-5	....PPGVL
mumps	....PSGLE
ndv	....VTTLF
rpv	LSAIDPTLD
measles	LSTDDPVID
pdv	IPVAHPSIE
cdv	IPVAHPSMK

Table F: Loop β3L23

bpi-iii	E.HEENGDVICNTTGCPGKTQRDC.....
sendai	T.TPLQGDTKCRTQGCQQVSQDTC.....
hpi-ii	ISGTPSYNKQSSRYFIPKHPNITCAGNSSEQAA.
sim-v-5	IKGTSLWNNQANKYFIPQMVAALCSQNQATQVQ.
mumps	LPNSTLGVKLAREFFRPVNPYNPCSGPQQDLDQ.
ndv	KPNSPSDTVQEGKYVIYKRYNDTCPDEQDYQIR.
rpv	Q....VKMGKCRLEACRDRPPPFCNSTDWEPLEA
measles	K....LRMETCFQQACKGKIQALCENPEWAPLKD
pdv	Q....GEQINCLRSACKRRTYPMCNQTSWEPFGD
cdv	Q....EEQKGCLESACQRKTYPMCNQASWEPFGG

Table G: Loop β4L01

bpi-iii	KGIDTTFSLRVWTIPMRQNYWG
sendai	DYLSERPKIRVTTIPITQNYLG
hpi-ii	LSDMHTARCNLVMFNNSQVMMG
sim-v-5	LRQDLTNECLVLPFSNDQVLMG
mumps	WNQILVTNCELVVPSNNQTLMG
ndv	VSTSLGEDPVLTVPPNTVTLMG
rpv	.....KVDIIS
measles	.....KIKIAS
pdv	.....SINVSV
cdv	.....QLNISF

Table H: Loop β4L23

bpi-iii	STSWHSKL
sendai	SSGWHSQL
hpi-ii	SSSWWSAS
sim-v-5	SNSWWPMT
mumps	STSWWPYE
ndv	GSSYFSPA
rpv	GTKYWLTT
measles	NNVYWLTI
pdv	LNSGWLTI
cdv	LNSGWLTI

Table I: Loop β5L01

bpi-iii	RINWTWHNVLSRPGNDECPWGHSCPDGCITGVYT.....DA
sendai	TINWTPHEALSRPGNKECNWYNKCPKECISGVYT.....DA
hpi-ii	EAQWVPSYQVPRPGVMPCNATSFCPANCITGVYA.....DV
sim-v-5	SAQNVPTQQVPRPGTGDCSATNRCPGFCLTGVYA.....DA
mumps	NMSWIPIYSFTRPGSGKCSGENVCPIACVSGVYL.....DP
ndv	LHSPYTFNAFTRPGSIPCQASARCPNSCVTGVYT.....DP
rpv	PSLKISPNILTLPIRSG.....GGDCYTPTYLSDRADDDVKLSS
measles	PRFKVSPYLFNVPIKEA.....GEDCHAPTYLPAEVDGDVKLSS
pdv	DQFIVTPHILTFAPRES.....STDCHLPIQTYQIQDDDVLLES
cdv	DQFTVLPHVLTFAPRES.....SGNCYLPIQTSQIRDRDVLIES

Table J: Loop β5L23

bpi-iii	....LDSQKSREN
sendai	....LYANTSRVN
hpi-ii	F...LRNESNRTN
sim-v-5	Y...LNTATQRIN
mumps	L...LNSSTTRVN
ndv	FGTMLDSEQARLN
rpv	.....TYDISRVE
measles	.....TYDTSRVE
pdv	.....TYDVSRSD
cdv	.....TYDISRSD

Table K: Loop β6L01

bpi-iii	YNRTLPAAY
sendai	KDVQLEAAY
hpi-ii	NNTNHKAAY
sim-v-5	GSSGQEAAY
mumps	GTQGLSASY
ndv	SSSSTKAAY
rpv	PVK.LPIKGDPVS
measles	PFR.LPIKGVPIE
pdv	PFR.LKTKGRPDI
cdv	PFR.LTTKGRPDF

Table L: Loop β6L23

bpi-iii	LNTFQPM					
sendai	LNTLQPM					
hpi-ii	LGEFQII					
sim-v-5	LGQFQIV					
mumps	VGEFQIL					
ndv	FGEFRIV					
rpv	T	RKQVTHT	and	T.	RKQVTHT	
measles	S	GGHITHS	and	T.	RKQVTHT	
pdv	T	NSTSVVE	and	T.	RKQVTHT	
cdv	A	NSTTSVE	and	T.	RKQVTHT	

Table M Similarities between neuraminidases and trans-neuraminodases of bacterial, protozoan, viral or eukaryotic origin.

	V.cholerae	A.viscosus	Tryp.cuzi	Salm.typh.	Clos.sept.	Rat.cytosolic	Influenza A	Influenza B	pI II	NDV	Sendai	pI III	CDV	Measles
V.cholerae		59	53	61	77	54	34	29	30	30	30	33	15	21
A.viscosus	59		48	69	77	51	33	28	25	26	30	28	26	20
Tryp.cruzi	53	48		78	59	48	31	37	26	31	32	34	30	26
Salm.typh.	61	69	78		94	56	30	31	33	33	33	27	25	27
Clos.sept.	77	77	59	94		64	42	29	30	23	32	31	29	24
Rat.cytosolic	54	51	48	56	64		23	25	31	31	21	25	25	24
Influenza A	34	33	31	30	42	23		112	35	32	31	27	24	25
Influenza B	29	28	37	31	29	25	112		38	29	28	27	25	23
pI II	30	25	26	33	30	31	35	38		151	105	107	38	36
NDV	30	26	31	33	23	31	32	29	151		114	103	33	41
Sendai	30	30	32	33	32	21	31	28	105	114		233	46	44
pI III	33	28	34	27	31	25	27	27	107	103	233		50	51
CDV	15	26	30	25	29	25	24	25	38	33	46	50		135
Measles	21	20	26	27	24	24	25	23	36	41	44	51	135	

Each comparison gives number of identical residues according to alignment of study.

CLAIMS

1. An isolated or recombinant proteinaceous substance comprising at least one virus epitope derived from an attachment protein of a virus from the family of paramyxoviridae, said epitope corresponding to an antigenic site present on the HN protein of paramyxovirus, which site is identified as one of loop β1L01, β1L23, β2L01, β2L23, β3L01, β3L23, β4L01, β4L23, β5L01, β5L23, β6L01 and β6L23, or a functional equivalent thereof.

2. An antibody specifically directed against a virus epitope derived from an attachment protein of a virus from the family of paramyxoviridae, said epitope corresponding to an antigenic site present on the HN protein of paramyxovirus, which site is identified as one of loop β1L01, β1L23, β2L01, β2L23, β3L01, β3L23, β4L01, β4L23, β5L01, β5L23, β6L01 and β6L23, or a functional equivalent thereof.

3. A method for selecting and producing an epitope of a first virus of the family of the paramyxoviridae, comprising aligning the sequence of the HN protein of said virus with the sequence of a second virus of the same family of which the 3-D structure is given, identifying the sequence of said first virus which corresponds with an epitope of said second virus and synthesizing or isolating a proteinaceous substance having said sequence or a functional equivalent thereof.

4. A method according to claim 3 wherein said second virus is a paramyxovirus.

5. A method according to claim 3 or 4 wherein said first virus is the bPIV-3 virus.

6. A vaccine composition comprising a proteinaceous substance according to claim 1.

7. A vaccine composition comprising a virus from the family of paramyxoviridae which virus is modified by functionally removing an immunodominant epitope, which immunodominant

epitope corresponds to one of the following sites which are identified as one of loop β1L01, β1L23, β2L01, β2L23, β3L01, β3L23, β4L01, β4L23, β5L01, β5L23, β6L01 and β6L23.

8. A vaccine composition comprising a virus from the family of paramyxoviridae which virus is modified by functionally removing an immunodominant epitope, which immunodominant epitope is selected by a method according to anyone of claims 3-5.

9. A diagnostic test comprising a substance according to claim 1 and/or an antibody according to claim 2.

10. A carbohydrate substance blocking enzymatic activity of a morbillivirus H protein, or a functional equivalent thereof.

11. A substance according to claim 10 which is a sialic acid.

12. A substance according to claim 11 which is a sialic acid modified at the 5 or 6 position.

13. A pharmaceutical composition comprising a substance according to anyone of claim 10-12 and a pharmaceutically acceptable carrier.